BU-147-M

TESTS FOR NON-ADDITIVITY

by

W. T. Federer, D. S. Robson, and John W. Tukey

1. INTRODUCTION

The heuristic procedure of partitioning the total sum of squares into meaningful components and an unexplained residual has proven to be one of the most popular means of analyzing data. Carried to completion in the form of single degree of freedom sums of squares, this process may lead to a full explanation of the variability in the data, as is commonly achieved in a two-level factorial system. The two-level complete factorial may also be singled out, however, as the only case for which there exists a unique natural or standard partition into individual degrees of freedom; in any other situation the standard partition is at best incomplete, and judgment must then be exercised in isolating any further degrees of freedom. Although this judgment will be based on the situation at hand, some general suggestions can be made concerning further possibilities associated with the standard layouts. These suggestions entail first the full exploitation of the method of fitting constants, followed by tests for non-linear associations employing the principles of Tukey's "one degree of freedom for non-additivity" [1949].

2. THE METHOD OF FITTING CONSTANTS

Applied to unbalanced layouts, the method of fitting constants has long been regarded with some distaste in everyday practice because of the tedious computations required for the solution of the normal equations. When an additional factor is observed (as a covariate, for example) in an otherwise standard balanced experiment there is a noticeable tendency to oversimplify the role of this factor in the model in order to retain the neat computational form characteristic of a balanced layout. With the now general availability and utilization of electronic computers, this pernicious subterfuge must be regarded with increasing alarm; simplifying assumptions based only on the grounds of computational convenience form an ever weaker crutch to the statistician.

The standard randomized blocks covariance analysis illustrates one of the more extreme forms of abuse to which we routinely subject our clients. A Model I randomized blocks covariance layout is, in effect, a three-factor design, the blocks or replicates representing levels of a composite environmental factor, the treatments representing levels of the second factor, and the covariate representing levels of an additional, third factor. The covariate factor is commonly assumed to have a linear effect and all interactions among the three factors are assumed away in the standard analysis, with no provisions for testing these sweeping assumptions of additivity.

By way of contrast, for the one-way covariance layout, or covariance in a completely randomized design, a test for interaction between the treatment and covariate factor is commonly recommended in the form of an F-test of homogeneity of within-treatment regression coefficients. Since an analogous interaction test in the two-way case may be constructed by the familiar method of fitting constants, our inconsistency in failing to routinely recommend the test in this case, even when the analysis is performed on a high speed computer, can only be ascribed to the archaic view that the method of fitting constants is too onerous for everyday use. A detailed description of this interaction test by the method of fitting constants is given in section 5.

3. TESTS OF NONLINEAR HYPOTHESES

The residual sum of squares remaining after the fitting of a linear model may be further partitioned by Tukey's procedure for isolating single degrees of freedom to test nonadditivity. This approach provides tests against nonlinear alternative models in which the effects of qualitative factors are functionally related to those of other qualitative or quantitative factors, and judgment based on subject matter knowledge may be exercised in choosing the particular functional relations to be tested. For example, if the alternative to the additive model

$$\text{ave}(Y_{ij}) = \mu + \rho_i + \gamma_j\ , \quad i = 1, \cdots, r;\ j = 1, \cdots, c$$

for a two-way classification is expressible as

$$\text{ave}(Y_{ij}) = \mu + \rho_i + \gamma_j + \theta f(\rho_i, \gamma_j),$$

where the functional form of f is specified, then the F statistic computed by applying the coefficients $c_{ij} = \dfrac{f_{ij}}{\sqrt{\Sigma\Sigma(f_{ij} - \bar{f}_{i.} - \bar{f}_{.j} + \bar{f}_{..})^2}}$, $f_{ij} = f(\bar{y}_{i.} - \bar{y}_{..}, \bar{y}_{.j} - \bar{y}_{..})$ $= f(R_i, C_j)$

$$c_{ij} = \frac{f(\bar{y}_{i.} - \bar{y}_{..}, \bar{y}_{.j} - \bar{y}_{..})}{\sqrt{\sum_{i=1}^{r}\sum_{j=1}^{c} f^2\cancel{(\bar{y}_{i.} - \bar{y}_{..}, \bar{y}_{.j} - \bar{y}_{..})}}} = \frac{f(R_i, C_j)}{\sqrt{\sum_1^r\sum_1^c f^2(R_i, C_j)}}$$

to the residuals

$$e_{ij} = Y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$$

as

$$F = \frac{\left(\sum_1^r\sum_1^c c_{ij}e_{ij}\right)^2[(r-1)(c-1) - 1]}{\sum_1^r\sum_1^c e_{ij}^2 - \left(\sum_1^r\sum_1^c c_{ij}e_{ij}\right)^2}$$

is distributed as Snedecor's F on 1 and (r-1)(c-1) - 1 degrees of freedom under the additive model with NIID (normal, independent and identically distributed) errors. The basis of this test is the independence between the least squares residuals and the estimators of the constants in a linear model.

In the original non-additivity test proposed by Tukey the functional relation of factor effects was taken as

$$f(\rho_i, \gamma_j) = \rho_i\gamma_j$$

providing a test against a multiplicative type of interaction effect. In fact, if the errorless model

$$Y_{ij} = \text{ave}(Y_{ij}) = \mu + \rho_i + \gamma_j + \theta\rho_i\gamma_j$$

holds then

$$\left(\sum_1^r\sum_1^c c_{ij}e_{ij}\right)^2 = \sum_1^r\sum_1^c e_{ij}^2$$

so in this case the residual from the additive model is completely accounted for in Tukey's one degree of freedom for non-additivity.

Federer, in his discussion of non-additivity [1955, pp. 49-51, 206-209], suggested another test of this nature based on the coefficient of regression of the yield of one treatment on that of another in a randomized blocks experiment. If the additive model holds then the true regression coefficient would be unity, and the alternative is a multiplicative type of interaction between treatments and blocks. Mandel [1961] formalized this approach by regressing a treatment yield on the block mean rather than the yield of another treatment, thus providing a perfect fit to an errorless non-linear model of the form

$$Y_{ij} = \text{ave}(Y_{ij}) = \mu + \rho_i + \gamma_j + P_i\gamma_j \; .$$

ρ_i ?

P_i not defined anywhere? $P_i = \rho_i\theta$?

By symmetry considerations, this approach is readily extended to a method for fitting the model

$$\text{ave}(Y_{ij}) = \mu + \rho_i^{(0)} + \gamma_j^{(0)} + \rho_i^{(1)}\gamma_j^{(0)} + \rho_i^{(0)}\gamma_j^{(1)} + \theta\rho_i^{(0)}\gamma_j^{(0)}$$

and in the <u>vacuum cleaner</u> method, Tukey [1962, pp. 49-60] presents a procedure for fitting the model

$$\text{ave}(Y_{ij}) = \mu + \rho_i^{(0)} + \gamma_j^{(0)} + \sum_{v=1}^{s}\Big(\rho_i^{(v)}\gamma_j^{(v-1)} + \rho_i^{(v-1)}\gamma_j^{(v)}$$
$$+ \theta^{(v)}\rho_i^{(v-1)}\gamma_j^{(v-1)}\Big)$$

using the restrictions

$$\sum_{i=1}^{c} \rho_i^{(v)} = \sum_{i=1}^{c} \rho_i^{(v)} \rho_i^{(v-t)} = 0 \ , \quad t = 1, \cdots, v$$

$$\sum_{j=1}^{r} \gamma_j^{(v)} = \sum_{j=1}^{r} \gamma_j^{(v)} \gamma_j^{(v-t)} = 0 \ , \quad t = 1, \cdots, v$$

Here $s \leq \min((r-1),(c-1))$, and in the case of equality the method provides a perfect fit to a two-way array of data. The details of this procedure are described in section 4.

4. NON-ADDITIVITY IN THE TWO-WAY CLASSIFICATION WITH ONE OBSERVATION PER CELL

Interaction of unrestricted form in the model

$$Y_{ij} = \mu + \rho_i + \gamma_j + (\rho\gamma)_{ij} + \epsilon_{ij} \ , \quad i = 1, \cdots, r; \ j = 1, \cdots, c$$

is completely confounded with error, and only some restricted type of interaction described by fewer than $(r-1)(c-1)$ parameters is subject to test. As already mentioned, a single-parameter interaction of the form $(\rho\gamma)_{ij} = \theta\rho_i\gamma_j$ (as would result, for example, if the additive scale were log y instead of y) is tested by Tukey's one degree of freedom for non-additivity,

$$b^2_{..} \sum_1^r R_i^2 \sum_1^c C_j^2 = \frac{\left(\sum_1^r \sum_1^c R_i C_j e_{ij}\right)^2}{\sum_1^r R_i^2 \sum_1^c C_j^2}$$

Mandel constructs a test for an (r-1)-parameter interaction of the form $(\rho\gamma)_{ij} = P_i\gamma_j, \ \sum_1^r P_i = 0$, as

? ?

not defined

Walt -

I've given further thought to the idea of expanding the "no-interaction model"

$$\mu_{ij} = f(\rho_i, \gamma_j)$$

in a Taylor series, in order to test for "interaction", and while I still think the idea has merit I don't feel that I can incorporate it into the present paper in the (negative) amount of time available. Some additional labor is required to work out the details of the procedure and the implications of a perfect fit to a finite number of terms from the Taylor series expansion; there may even be a thesis in these details.

The error on page 3 (which I had noted earlier but couldn't recall) is now indicated on your copy.

$$F_{(r-1),(r-1)(c-2)} = \frac{(c-2) \sum_1^r b_{i*}^2 \sum_1^c C_j^2}{\sum_1^r \sum_1^c e_{ij}^2 - \sum_1^r b_{i*} \sum_1^c C_j^2}$$

where b_{i*} is the coefficient of regression of $Y_{ij} - \bar{y}_{.j}$ on $\bar{y}_{.j}$,

$$b_{i*} = \frac{\sum_1^c C_j e_{ij}}{\sum_1^c C_j^2} , \quad \sum_1^r b_{i*} = 0 .$$

As noted by Mandel, Tukey's one degree of freedom can be partitioned out of the sum of squares among the b_{i*}, implying that $b_{..}$ is a linear function of the b_{i*} . We note in fact that

$$b_{..} = \frac{\sum_1^r b_{i*} R_i}{\sum_1^r R_i^2}$$

and, defining

$$b_{i.} = b_{i*} - b_{..}R_i , \quad \sum_1^r b_{i.} = \sum_1^r b_{i.}R_i = 0$$

then

$$\sum_1^r b_{i*}^2 \sum_1^c C_j^2 = \sum_1^r b_{i.}^2 \sum_1^c C_j^2 + b_{..}^2 \sum_1^r R_i^2 \sum_1^c C_j^2$$

Tukey's $b_{..}$ is independent of $\{b_{i.}\}$ and, by symmetry, independent of $\{b_{.j} = b_{*j} - b_{..}C_j\}$, where

$$b_{*j} = \frac{\sum_1^r R_i e_{ij}}{\sum_1^r R_i^2}$$

and, as easily verified, $\{b_{i.}\}$ is independent of $\{b_{.j}\}$. Since the new residual

$$d_{ij} = e_{ij} - b_{..}R_iC_j - b_{i.}C_j - b_{.j}R_i = e_{ij} - b_{i*}C_j - b_{*j}R_i + b_{..}R_iC_j$$

is, for fixed $\{R_i\}$ and $\{C_j\}$, independent of b, $\{b_{i.}\}$, $\{b_{.j}\}$, we conclude that

$$F_{(r+c-s),(r-2)(c-2)} = \frac{(r-2)(c-2)\left[\sum_1^r b_{i*}^2 \sum_1^c C_j^2 + \sum_1^c b_{*j}^2 \sum_1^r R_i^2 - b_{..}^2 \sum_1^r R_i^2 \sum_1^c C_j^2\right]}{(r+c-3)\left[\sum_1^r \sum_1^c e_{ij}^2 - \sum_1^r b_{i*}^2 \sum_1^c C_j^2 - \sum_1^c b_{*j}^2 \sum_1^r R_i^2 + b_{..}^2 \sum_1^r R_i^2 \sum_1^c C_j^2\right]}$$

provides a test for interaction of the form $(\rho\gamma)_{ij} = P_i\gamma_j + \rho_i\Gamma_j + \theta\rho_i\gamma_j$. This type of interaction would arise, for example, if

$$y_{ij} = (\mu + p_i + q_j)a^{(\mu+p_i+q_j)} .$$

μ ?

$F_j \neq \theta\gamma_j$

$P_i \neq \theta\rho_i$

In this case, or whenever the errorless model

$$y_{ij} = \mu + \rho_i + \gamma_j + P_i\gamma_j + \rho_i\Gamma_j + \theta\rho_i\gamma_j$$

holds then the new residual d_{ij} vanishes,

$$d_{ij} = e_{ij} - b_{..}R_iC_j - b_{i.}C_j - b_{.j}R_i = 0 \;; \quad i = 1, \cdots, r; \; j = 1, \cdots, c .$$

As Tukey [1962, p. 53] points out, this partitioning procedure may be continued in various ways until the (r-1)(c-1) degrees of freedom in error are exhausted. As the next step we may define

$$B_{..} = \frac{\sum_1^r \sum_1^c b_{i.} b_{.j} d_{ij}}{\sum_1^r b_{i.}^2 \sum_1^c b_{.j}^2} \;;$$

$$B_{i*} = \frac{\sum_1^r b_{.j} d_{ij}}{\sum_1^c b_{.j}^2} , \quad \sum_1^r B_{i*} = \sum_1^r B_{i*} R_i = 0$$

$$B_{i.} = B_{i*} - B_{..}b_{i.}\ ,\ \sum_1^r B_{i.} = \sum_1^r B_{i.}R_i = \sum_1^r B_{i.}b_{i.} = 0$$

and, similarly,

$$B_{.j} = B_{*j} - B_{..}b_{.j}\ ,\ \sum_1^c B_{.j} = \sum_1^c B_{.j}C_j = \sum_1^c B_{.j}b_{.j} = 0$$

where now for fixed $\{R_i\}$ and $\{C_j\}$, the sets of statistics $b_{..}$, $\{b_{i.}\}$, $\{b_{.j}\}$, $B_{..}$, $\{B_{i.}\}$, $\{B_{.j}\}$ with $1 + (r-2) + (c-2) + 1 + (r-3) + (c-3) = 2r + 2c - 8$ degrees of freedom are mutually independent and are independent of the new residuals $\{f_{ij}\}$,

$$f_{ij} = d_{ij} - B_{i*}b_{.j} - B_{*j}b_{i.} + B_{..}b_{i.}b_{.j}$$

with $(r-3)(c-3)$ degrees of freedom. The corresponding F-statistic with $(2r+2c-8)$ and $(r-3)(c-3)$ degrees of freedom therefore provides a test for interaction of the form $(\rho\gamma)_{ij} = P_i\gamma_j + \rho_i\Gamma_j + \theta\rho_i\gamma_j + \beta_{i.}\Gamma_j + \beta_{.j}P_i + \beta_{..}P_i\Gamma_j$ which arises, for example, when

$$y_{ij} = (\mu + p_i + q_j)^2 a^{(\mu+p_i+q_j)}$$

5. NON-ADDITIVITY IN THE ANALYSIS OF COVARIANCE

When an additional factor in the form of a covariate is included in an otherwise balanced design the data are most commonly analyzed as though the effect of this concomitant factor were linear and additive with the treatment and design factors. Thus the simplest type of non-additivity to be considered is heterogeneity of slope of the linear regressions at different levels of the balanced factors. A test for this sort of interaction between the concomitant factor and a single treatment or design factor follows directly from the method of fitting constants though, following the approach used by Robson and Atkinson [1960] in testing additivity in a one-way covariance analysis, the method of fitting constants may be combined with Tukey's approach to obtain tests more sensitive against specific non-linear alternatives to the additive model.

To illustrate this combined approach we consider a balanced three-factor experiment, supplemented with a covariate, in which two of the factors, say A and B, denote the controlled treatment factors and the third denotes the design factor-blocks, or replicates. The usual covariance Model I is then

$$\underline{\text{additive:}}\quad Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \rho_k + \beta X_{ijk} + \epsilon_{ijk} \quad \begin{cases} i = 1, \cdots, a \\ j = 1, \cdots, b \\ k = 1, \cdots, r \end{cases}$$

where $\{\rho_k\}$ are the replicate effects and $\{\beta X_{ijk}\}$ are the additive effects of the concomitant factor at the levels $\{X_{ijk}\}$. A non-additive alternative model in which A interacts with the linear effect of the concomitant factor is obtained by attaching an i-subscript to the slope β or, in keeping with the structure of the rest of the model, by adding a term $\beta_i X_{ijk}$

$$\underline{\text{non-additive:}}\quad Y_{ijk} = \mu + \alpha_i + \tau_j + (\alpha\tau)_{ij} + \rho_k + \beta X_{ijk} + \beta_i X_{ijk} + \delta_{ijk}$$

where $\{\beta_i\}$ are subject to a linear constraint, such as $\beta_a = 0$.

The residuals in the completely additive model

$$\begin{aligned} e_{ijk} &= Y_{ijk} - \bar{y}_{ij.} - \bar{y}_{..k} + \bar{y}_{...} - b_{...}(X_{ijk} - \bar{x}_{ij.} - \bar{x}_{..k} + \bar{x}_{...}) \\ &\equiv y_{ijk} - b_{...}x_{ijk} \end{aligned}$$

where

$$b_{...} = \frac{\sum_{ijk} x_{ijk} y_{ijk}}{\sum_{ijk} x^2_{ijk}}$$

are independent of $b_{...}$ and the estimators of the other constants in the linear model if the errors $\{\epsilon_{ijk}\}$ are NIID. Since the $\{e_{ijk}\}$ satisfy the constraints

$$\sum_k e_{ijk} = \sum_{ij} e_{ijk} = \sum_{ijk} e_{ijk} X_{ijk} = 0$$

then the residual sum of squares has $(ab-1)(r-1) - 1$ degrees of freedom. When the interaction term $\beta_i X_{ijk}$ is appended to the model the balance is lost, and the new residuals $\{d_{ijk} = e_{ijk} - \hat{\beta}_{i..} x_{ijk}\}$ can be obtained only by fitting the constants $\{\beta_i\}$ to the additive residuals $\{e_{ijk}\}$.

As in any multiple regression problem there is some interest in examining and perhaps testing the individual linear regression coefficients before undertaking the matrix inversion required to compute the multiple regression coefficients. For the i^{th} level of factor A the regression coefficient $b_{i..}$ is given by

$$b_{i..} = \frac{\sum_{jk} X_{ijk} e_{ijk}}{\sum_{jk} (X_{ijk} - \bar{x}_{i..})^2}$$

and, for testing purposes, the variance of the numerator is $\sigma_\epsilon^2 V_{ii}$,

$$V_{ii} = \sum_{jk} X_{ijk}^2 - \frac{1}{r} \sum_{j}(\sum_{k} X_{ijk})^2 - \frac{1}{a}\left[\frac{1}{b}\sum_{k}(\sum_{j} X_{ijk})^2 - \frac{1}{rb}(\sum_{jk} X_{ijk})^2\right] - \frac{(\sum_{jk} X_{ijk} x_{ijk})^2}{\sum_{ijk} X_{ijk} x_{ijk}} .$$

The squared numerator divided by V_{ii} may then be tested against the residual

$$\sum_{ijk} e_{ijk}^2 - \frac{(\sum_{jk} X_{ijk} e_{ijk})^2}{V_{ii}}$$

with $(ab-1)(r-1) - 2$ degrees of freedom.

The multiple regression coefficients $(\hat{\beta}_{i..})' = (\hat{\beta}_{1..}, \cdots, \hat{\beta}_{a-1..})$ are obtained by inverting the covariance matrix (V_{ih}), $i, h = 1, \cdots, a - 1$ where, for $i \neq h$,

$$V_{ih} = -\frac{1}{a}\left[\frac{1}{b}\sum_{k}(\sum_{j} X_{ijk})(\sum_{j} X_{hjk}) - \frac{1}{rb}(\sum_{jk} X_{ijk})(\sum_{jk} X_{hjk})\right] - \frac{\sum_{jk} X_{ijk} x_{ijk} \sum_{jk} X_{hjk} x_{hjk}}{\sum_{ijk} X_{ijk} x_{ijk}}$$

and then

$$(\hat{\beta}_{i..}) = (V_{ih})^{-1}(\sum_{jk} X_{1jk}e_{1jk}, \cdots, \sum_{jk} X_{a-1\ jk}e_{a-1\ jk})'$$

or, more briefly,

$$\hat{\beta}_{(A)} = V_{(A)}^{-1}X'_{(A)}e \ .$$

The sum of squares with a-1 degrees of freedom due to the $\{\beta_{i..}\}$,

$$\text{sum of squares} = e'X_{(A)}V_{(A)}^{-1}X'_{(A)}e = \hat{\beta}'_{(A)}V_{(A)}\hat{\beta}_{(A)}$$

may then be tested against the new residual

$$\sum_{ijk} d_{ijk}^2 = \sum_{ijk} (e_{ijk} - \hat{\beta}_{i..}x_{ijk})^2 = e'e - \hat{\beta}'_{(A)}V_{(A)}\hat{\beta}_{(A)}$$

with (ab-1)(r-1) - (a-1) degrees of freedom.

A single degree of freedom analogous to Tukey's one degree of freedom for non-additivity may be partitioned out of this sum of squares due to $\{\beta_{i..}\}$ to provide a test against the alternative hypothesis that the slopes $\beta_{i..}$ are proportional to the additive effects α_i. Under this particular alternative the regression lines at the different levels of factor A form a pencil, intersecting at a common point, in contrast to the family of parallel lines specified by the additive model. A pencil with this property is characterized by a linear relation between slope and intercept, and is tested from the regression of treatment effect

$$A_i = \bar{y}_{i..} - \bar{y}_{...} - b_{...}(\bar{x}_{i..} - \bar{x}_{...})$$

on slope $\hat{\beta}_{i..}$. Thus, denoting by $(A_i)'$ the row vector $(A_1, \cdots, A_{a-1})$ we obtain the test

$$F_{1,(ab-1)(r-1)-(a-1)} = \frac{[(A_i)'(\hat{\beta}_{i..})]^2[(ab-1)(r-1) - (a-1)]}{(A_i)'V_{(A)}^{-1}(A_i) \sum_{ijk} d_{ijk}^2}$$

which combines the two approaches for testing non-additivity.

6. HETEROGENEITY OF ERROR VARIANCE IN A TWO-WAY CLASSIFICATION

Empirical evidence from agricultural experiments indicates that yield data commonly depart from the analysis of variance model by exhibiting heterogeneous error variance, usually in the form of a monotone relation between mean and variance. A test of homogeneity having sensitivity against this patterned alternative may be constructed by comparing the residual mean squares associated with the treatments giving the lowest and highest mean yields. Since residuals are statistically independent of estimated treatment effects under the homogeneity assumption, this selection of residual mean squares according to the rank of the treatment effect then has no influence on the distribution of the selected mean square ratio. In a one-way classification, for example, the ratio of any two within-class mean squares is distributed as F under the usual assumptions, regardless of the rank of these two class means. Under the alternative model with unequal error variances related to the treatment means, however, the selection of residuals for comparison according to treatment does influence the distribution of the resulting mean square ratio, which then becomes a mixture of non-central F-distributions.

In the randomized blocks case the residual mean squares associated with the different treatments are not statistically independent, and the comparison of two such mean squares is not directly achievable in the form of an F-test. A simple orthogonal transformation, however, produces the desired result of two independent mean squares whose difference in average value under the alternative hypothesis is a linear function of the true difference in error variance. Thus, instead of attempting to compare the two dependent residual mean squares E_1 and E_2,

$$E_1 = \sum_{j=1}^{r} e_{1j}^2/(r-1)$$

where treatments 1 and 2 may have been selected for having the lowest and highest means, respectively, we compare the independent mean squares

$$E_1' = \frac{c}{(r-1)(c-1)} \sum_{j=1}^{r} e_{1j}^2$$

$$E_2' = \frac{c-1}{(r-1)(c-2)} \sum_{j=1}^{r} \{\frac{e_{1j}}{c-1} + e_{2j}\}^2$$

each with r-1 degrees of freedom.

Pooling residuals for several low ranking and for several high ranking treatments may increase the sensitivity of this test procedure through the increase in degrees of freedom, and may be accomplished by an extension of the above transformation. In general, the transformed residual mean square for the k^{th} treatment,

$$E_k' = \frac{c-k+1}{(r-1)(c-k)} \sum_{j=1}^{r} \{\frac{\sum_{i=1}^{k-1} e_{ij}}{c-k+1} + e_{kj}\}^2$$

is independent of $E_1', \cdots, E_{k-1}'$ under the usual assumptions, and the pooled sum of squares for the first k treatments is then given by

$$S_k' = (r-1) \sum_{i=1}^{k} E_i' = \sum_{j=1}^{r} \sum_{i=1}^{k} e_{ij}^2 + \frac{1}{c-k} \sum_{j=1}^{r} \{ \sum_{i=1}^{k} e_{ij}\}^2$$

with $k(r-1)$ degrees of freedom $(k \leq c-1)$. Thus, for comparing the pooled residual mean squares for the lowest k_1 and highest $k_2 = k - k_1$ ranking treatments we have

$$F = \frac{k_1}{k_2} \{\frac{S_{k_1+k_2}' - S_{k_1}'}{S_{k_1}'}\}$$

which follows the central F-distribution with $k_2(r-1)$ numerator degrees of freedom and $k_1(r-1)$ denominator degrees of freedom under the hypothesis of additivity and NIID errors.

The ratio of mean square expectations under the alternative model with error variance σ_{ij}^2 in cell ij is

$$1 + \frac{\{1 - \frac{1}{c-k}\}(\sigma_2^2 - \sigma_1^2) + \frac{k}{(c-k_1)(c-k)}(\sigma^2 - \sigma_1^2)}{\sigma_1^2 + \frac{1}{c-k_1}(\sigma^2 - \sigma_1^2)}$$

where

$$\sigma_1^2 = \frac{1}{rk_1} \sum_{i=1}^{k_1} \sum_{j=1}^{r} \sigma_{ij}^2$$

$$\sigma_2^2 = \frac{1}{rk_2} \sum_{i=k_1+1}^{k} \sum_{j=1}^{r} \sigma_{ij}^2$$

$$\sigma^2 = \frac{1}{rc} \sum_{i=1}^{c} \sum_{j=1}^{r} \sigma_{ij}^2$$

Viewing this as a one-tailed F-test against the alternative hypothesis that error variance is an increasing function of the treatment mean we see that unless substantial ranking errors occur then $\sigma_2^2 \geq \sigma^2 \geq \sigma_1^2$ and the ratio of mean square expectations exceeds unity accordingly. As mentioned before, however, this excess cannot be construed as the non-centrality parameter since the non-central distribution is actually a mixture of non-central F-distributions under this ranking and selection procedure.

7. RANK ANALYSIS OF ORTHOGONAL RANDOMIZED BLOCK RESIDUALS

An orthogonal transformation of the rc correlated residuals e_{ij} of a randomized block into $(r-1)(c-1)$ uncorrelated and homoscedastic residuals E_{ij} permits the application of a number of other techniques of residual analysis. The half-normal plot, for example, is of questionable validity when applied to the correlated residuals e_{ij} which are subject to $r + c - 1$ linear constraints, and the modification of such techniques to account for the correlation would seemingly lead to forbidding numerical analysis problems in the calculation of exact significance levels. It would appear, also, that these calculations would depend explicitly upon the design matrix, thus losing the generality characteristic of techniques devised for orthogonal residuals.

Since the purpose of such general techniques usually includes the identification of aberrant observations in the original sample, one condition on the orthogonal transformation is that each transformed residual (which may be a linear function of all true residuals ϵ_{ij}) retain a strong association with an original observation. An aberrant E_{ij} should thus point to a particular cell (i',j') in the $r \times c$ table, and for simplicity it would be desirable to have $(i,j) = (i',j')$. This is not true, for example, of the orthogonal residuals generated in sections 3-5; but there, no such association was needed to satisfy the purposes of the analysis. Since there are only $(r-1)(c-1)$ orthogonal residuals, the unique association of E_{ij} with cell (i,j) in the $r \times c$ table can hold for at most $(r-1)(c-1)$ cells.

A set of $(r-1)(c-1)$ residuals $\{E_{ij}\}$ which largely satisfy the above conditions, which appear well suited for general purposes and especially well suited for the purpose of testing heteroscedasticity, is given by the Helmert partition of the row $\times$ column interaction sum of squares:

$$\left.\begin{array}{l}\text{row 1 vs. remaining (r-1) rows}\\ \text{row 2 vs. remaining (r-2) rows}\\ \qquad\vdots\\ \text{row r-1 vs. row r}\end{array}\right\} \times \left\{\begin{array}{l}\text{column 1 vs. remaining (c-1) cols.}\\ \text{column 2 vs. remaining (c-2) cols.}\\ \qquad\vdots\\ \text{column (c-1) vs. column c}\end{array}\right.$$

where rows and columns have been rearranged so that $\bar{Y}_{1.} \geq \cdots \geq \bar{Y}_{r.}$ and $\bar{Y}_{.1} \geq \cdots \geq \bar{Y}_{.c}$. The transformed residual E_{ij},

E^2_{ij} = SS(row i vs. remaining (r-i) rows × col. j vs. remaining (c-j) cols.),

then has the form

$$E_{ij} = e^*_{ij}\sqrt{\frac{(r-i+1)(c-j+1)}{(r-i)(c-j)}}$$

where e^*_{ij} is a residual calculated in the usual manner but from the partial two-way table consisting only of rows i, i+1, ···, r and columns j, j+1, ···, c. Thus, except for a scalar, E_{ij} is an ordinary type of residual and so retains an ordinary type of association with cell (i,j). An analogous transformation has been previously applied to regression residuals for similar purposes by Hedayat and Robson [1970].

Ranking the rows and columns according to their observed means has no effect upon the probability distribution of $\{E_{ij}\}$ when the true errors $\{\epsilon_{ij}\}$ are NIID. On the other hand, if the errors ϵ_{ij} are independent but heteroscedastic with variance σ^2_{ij} which is a monotonic function of ρ_i and τ_j then, provided the rows and columns are correctly ranked, the E_{ij} have variances (estimated by E^2_{ij}) which are correspondingly monotonic in i and j. Only monotonicity---not the specific functional relation---is preserved by this transformation; for example, if σ^2_{ij} is a linear function of $\mu + \rho_i + \tau_j$,

$$\sigma^2_{ij} = \alpha + \beta(\mu + \rho_i + \tau_j)$$

then

$$\mathrm{var}(E_{ij}) = \alpha + \beta\left[\left(\frac{r-i-1}{r-i}\right)\rho_i + \left(\frac{c-j-1}{c-j}\right)\tau_j\right]$$

Non-linear functional relations are altered even more drastically by $\mathrm{var}(E_{ij})$ (and also by $\mathrm{var}(e_{ij})$). For this reason, and because the functional form of a heteroscedastic alternative hypothesis is usually unspecified in practice, we turn now to consideration of rank tests.

The monotone heteroscedastic alternative hypothesis implies that every row of the $(r-1) \times (c-1)$ matrix (E^2_{ij}) should be monotone in the same direction, and likewise every column. In the matrix (r_{ij}) where

$$r_{ij} = \text{rank of } E^2_{ij} \text{ in the row } (E^2_{i1}, E^2_{i2}, \cdots, E^2_{i(c-1)})$$

the entries in the i^{th} row should thus be highly correlated with j, and this correlation should be similar for all rows. Likewise, the matrix (c_{ij}) of within-column ranks should display the corresponding property of c_{ij} being highly and similarly correlated with i in every column. The "null" hypothesis of NIID residuals ϵ_{ij}, on the other hand, implies that each row of (r_{ij}) is simply a random, independent permutation of the column numbers 1, 2, $\cdots$, c-1; and similarly the columns of (c_{ij}) are independent random permutations of 1, 2, $\cdots$, r-1 (though the permutations in (r_{ij}) are not statistically independent of those in (c_{ij})). In testing one hypothesis against the other we are thus led to consider rank correlations, or functions thereof, as test statistics.

Within a row of (r_{ij}) we may measure rank correlation by Spearman's Rho or, normalized to have unit variance,

$$R_i = (\text{Spearman's Rho}) \sqrt{c-2}$$

$$= \frac{12}{c(c-1)\sqrt{c-2}} \sum_{j=1}^{c-1} (j - \tfrac{c}{2}) r_{ij}$$

and the corresponding within-column statistic is

$$C_j = \frac{12}{r(r-1)\sqrt{r-2}} \sum_{i=1}^{r-1} (i - \tfrac{r}{2}) c_{ij} \ .$$

Each R_i and C_j could be tested separately using tabled critical values for Spearman's Rho or the large sample normal approximation, where a sample size of 15 (rows or columns) is "large". The tests in rows and columns would not be statistically independent, however, and the error rate associated with any combination of the tests would be unknown.

An alternative procedure, which avoids (asymptotically) the problems of dependence between tests, utilizes $\bar{R}$ and $\bar{C}$ to test monotonicity within rows and columns. Though not independent these means are uncorrelated, and the distributions of $\bar{R}\sqrt{r-1}$ and $\bar{C}\sqrt{c-1}$ rapidly approach the standard normal as r and c respectively get large; for example, for c-1 = 2 columns the exact distribution of $\bar{R}\sqrt{r-1}$ is that of a normalized binomial variable with $p = \frac{1}{2}$ and $n = r-1$, and for $c-1 > 2$ the symmetric distribution of $\bar{R}\sqrt{r-1}$ approaches normality faster than the symmetric binomial. The approach to normality is accompanied by a corresponding approach to independence between the two tests.

The monotone heteroscedastic model is further characterized by uniform monotonicity within rows and within columns, which suggests that homogeneity among the R_i and among the C_j should be tested. Since the R_i (normalized Spearman's Rho) are mutually independent and approximately normally distributed for moderate size c when the null hypothesis is true, then the corrected sum of squares of the R_i is approximately chi-square distributed on r-2 degrees of freedom, and approximately independent of $\bar{R}$ since $R_i - \bar{R}$ and $\bar{R}$ are uncorrelated. An analogous remark applies to the C_j; however, $R_i - \bar{R}$ and $C_j - \bar{C}$ are correlated,

$$\mathrm{Cov}(R_i - \bar{R},\, C_j - \bar{C}) = \frac{(r-1)(c-1)}{(r-2)(c-2)}\,(i - \tfrac{r}{2})(j - \tfrac{c}{2})\mathrm{Cov}(r_{11}, c_{11})$$

and in order to achieve the desired independence between row-homogeneity and column-homogeneity tests, it becomes necessary to sacrifice one degree of freedom from each. Since the above covariance is in the form of $(i - \frac{r}{2})(j - \frac{c}{2})$ multiplied by a constant, any two linear contrasts, $\Sigma a_i(R_i - \bar{R})$ and $\Sigma b_j(C_j - \bar{C})$ are seen to be uncorrelated if and only if $\Sigma a_i(i - \frac{r}{2}) = \Sigma b_j(j - \frac{c}{2}) = 0$. Thus, by making the additional correction to each sum of squares,

$$X_R^2 = \sum_1^{r-1} (R_i - \bar{R})^2 - \frac{12}{r(r-1)(r-2)} \left[\sum_1^{r-1} (i - \frac{r}{2}) R_i \right]^2$$

$$X_C^2 = \sum_1^{c-1} (C_j - \bar{C})^2 - \frac{12}{c(c-1)(c-2)} \left[\sum_1^{c-1} (j - \frac{c}{2}) C_j \right]^2$$

the effect of the correlation is eliminated and these two test statistics are distributed approximately as independent chi-squares on r-3 and c-3 degrees of freedom, respectively.

In summary, the steps in this procedure are:

1) Calculate row and column means of the original Y_{ij} table and rearrange rows and columns so the row means and column means are in decreasing order.

2) Partition the interaction sum of squares of the rearranged table by calculating the interaction mean square E_{ij}^2 in each of the 2 × 2 tables

	col(j)	col(j+1) + col(j+2) + ··· + col(c)
row(i)		
row(i+1) + ··· + row(r)		

for i = 1, ···, r-1 and j = 1, ···, c-1 .

3) Construct the (r-1) × (c-1) table of mean squares E_{ij}^2, retaining the row and column ordering arrived at in Step 2, and derive from this a table of rank order within rows (r_{ij}) and a table of rank order within columns(c_{ij}).

4) For each row of (r_{ij}) (column of (c_{ij})) calculate the normalized Spearman rank correlation with column (row) number, the mean of these correlations, and their sum of squares corrected for both the mean and for the linear effect of rows (columns).

The resulting chi-square test statistics, $(r-1)\bar{\bar{R}}^2$, X_R^2, $(c-1)\bar{\bar{C}}^2$, X_C^2 are then approximately distributed as independent chi-squares, and hence their sum is approximately chi-square on $r + c - 4$ degrees of freedom.

REFERENCES

Federer, W. T. (1955). *Experimental Design Theory and Application*. Macmillan, New York.

Hedayat, A. and Robson, D. S. (1970). Independent stepwise residuals for testing homoscedasticity. Accepted for publication in the *J. Amer. Stat. Assoc.*

Mandel, John (1961). Non-additivity in two-way analysis of variance. *J. Amer. Stat. Assoc.* 56:878-888.

Robson, D. S. and Atkinson, G. F. (1960). Individual degrees of freedom for testing homogeneity of regression coefficients in a one-way analysis of covariance. *Biometrics* 16:593-605.

Tukey, John W. (1949). One degree of freedom for non-additivity. *Biometrics* 5:232-242.

Tukey, John W. (1962). The future of data analysis. *Ann. Math. Stat.* 33:1-67.